TECHNISCHE UNIVERSITÄT DRESDEN
Center for Information Services and High Performance Computing (ZIH)

Past, present, future installations: lessons learned

Energy Efficient HPC WG Workshop, November 16th 2015

Daniel Hackenberg (daniel.hackenberg@tu-dresden.de)

ZIH
Center for Information Services & High Performance Computing

LZR: Data Center at TU Dresden

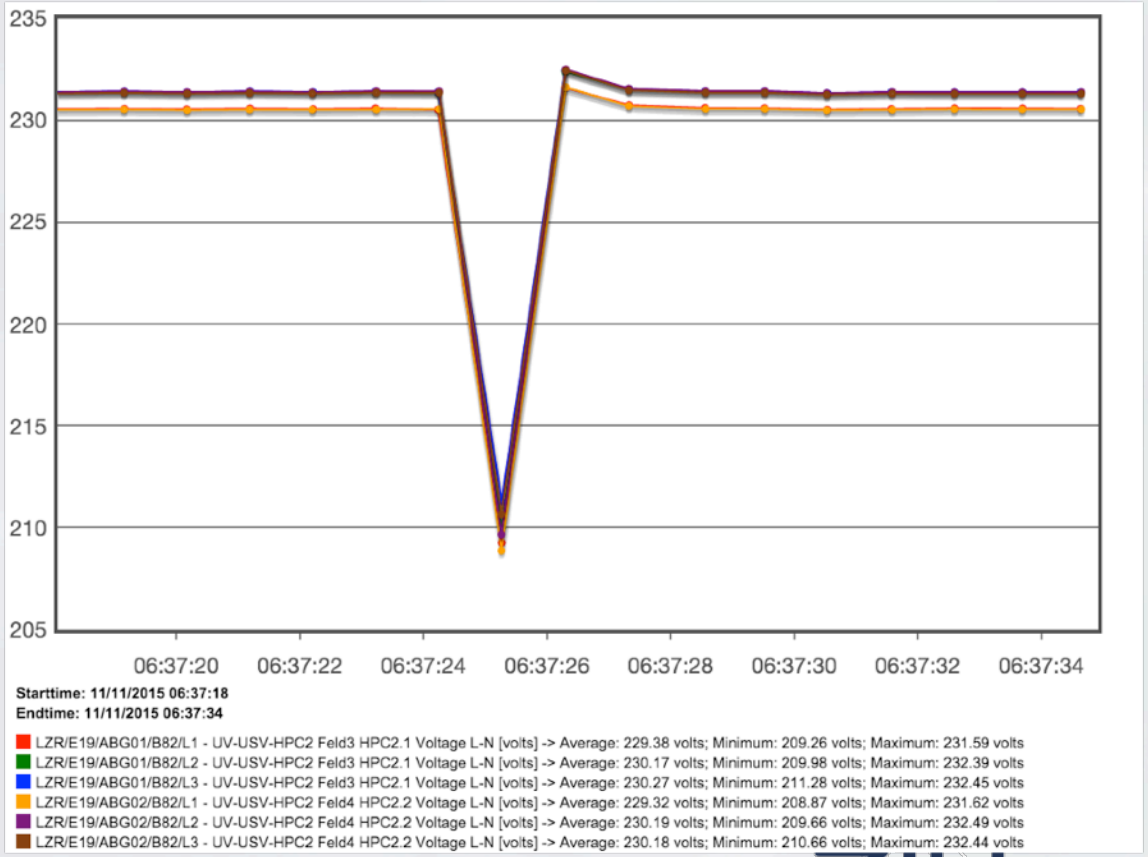
- Inauguration 07/2015
- Designed for ~5 MW IT power
- 3 cooling loops: 10°C, 15°C, 35°C
- Dedicated spaces for water-cooled high-density HPC, regular air-cooled IT, low-density networking and tapes
- Heat reuse (also in the summer)
- BACnet-based building automations system
  - Communications protocol for building automation and control networks
  - ANSI/ASHRAE norm 135, ISO norm 16484-5

Taurus HPC System at TU Dresden: Overview

- Taurus HPC system details
  - >1.3 PFLOPS from 43,664 Intel cores (>80% Haswell)
  - 300 TFLOPS from 128x NVIDIA K80
  - >108 TB RAM
  - 5 PB Lustre + >220 TB SSD
  - Water-cooled @ 35°C in, 45-50°C out
- HDEEM power monitoring infrastructure for 1500 Haswell nodes
  - 1000 samples/s from calibrated (2%) node power probes
  - 100 samples/s from VRs for 2xCPU and 4x DIMM

Infrastructure Lessons Learned

- Plenum concept works
- No matter how much data you save – at some point you will always have the need for more detail
- Water quality is wizardry; standards for water treatment would be helpful (instead of a lengthy list of requirements provided by the HPC vendor)



BAS Data Collection

- Data collection
  - Mostly BACnet, uses Python library BACpypes
  - ```
    $ ./ReadProperty.py
    > read 192.168.9.174 analogValue 19 presentValue
    20.2999992371
    > read 192.168.9.174 analogValue 19 units
    degreesCelsius
    ```
  - BACnet source queries multiple objects via ReadPropertyMultiple to
    - reduce BACnet traffic and
    - limit number of parallel request to devices that otherwise cannot handle the load (buffer overflows)
  - Only periodic polling, no change over value (CoV) used yet
  - HTTP source for Janitza power analyzers due to resolution limits
  - SNMP source for Piller UPS/Diesel
    ```

BAS Data Storage and Processing

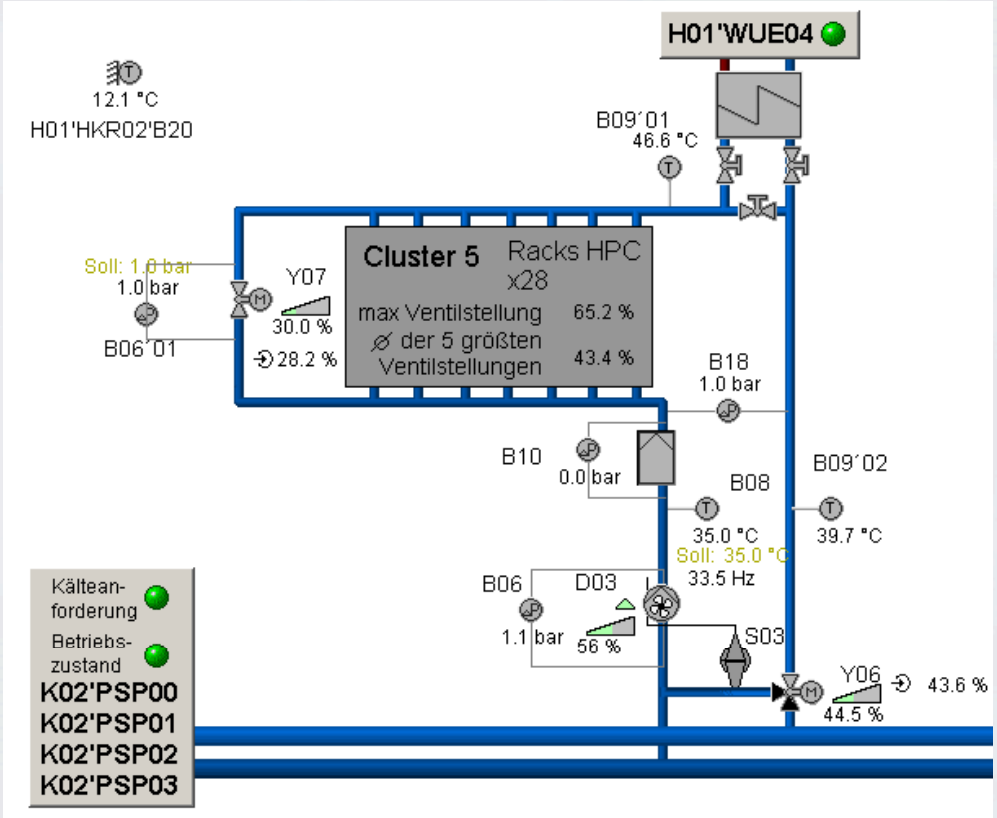
- Available data points at LZR

	# of devices	# of objects available	# of objects recorded	values [1/s]	Volume [GB/year]
Emerson CRAH units	26	9442	502	43	40
Janitza power analyzers	59	3599	216	170	160
Siemens BAS	26	13610	1223	69	65
Jaeggi cooling towers	5	295	15	1	1
Sum	116	26946	1956	283	266

- All data goes into Dataheap storage (via dhlib.py)
- Dataheap also stores data from thousands of sensors in the HPC machine

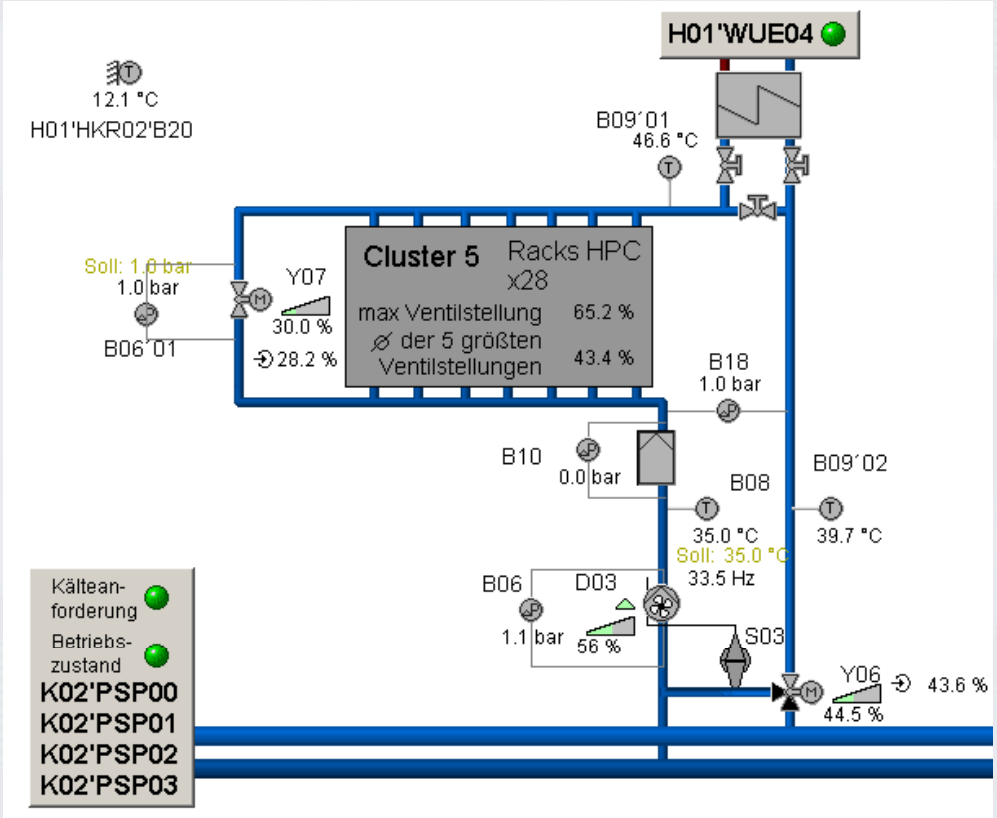
Feedback from Dataheap Siemens BAS

- Admin nodes collect data (e.g. from heat exchangers, PSUs) and send this to LZR-Dataheap
- BACnet-Server on LZR-Dataheap processes data for 3 cooling loops, each with
  - max. valve opening and
  - average of 5 largest valve openings
- BACnet-Server uses BACpypes to simulate BACnet device with 6 objects in the BAS network
- Siemens Desigo collects this data



Interaction Between Internal HPC Controls and BAS Controls

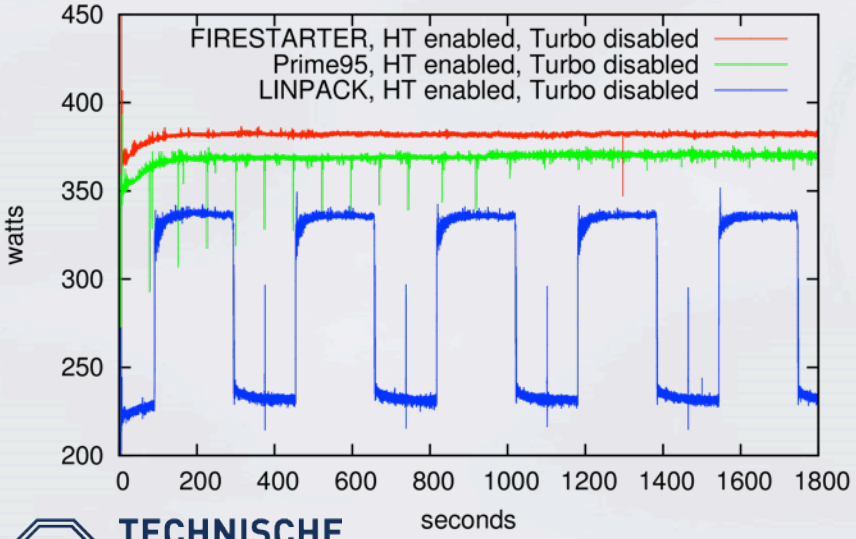
- Options to implement control loops
  - Control of overflow valve based on 5 largest valve openings to ensure high delta-T
  - Control of pump speed based on single largest valve opening to reduce pump power at partial load
- Issues
  - Guarantee availability of data points or
  - Detect errors and implement alternatives



FIRESTARTER: A Processor Stress Test Utility

- FIRESTARTER: Invaluable tool for all building control/load tests
  - Version 1.3 supports Intel (up to Broadwell-H), AMD (family 15h), Nvidia (CUDA)
  - http://tu-dresden.de/zih/firestarter/

Intel Xeon E5-2670, Sandy Bridge-EP (2P), AVX routine



Dual socket Intel Haswell system with two NVIDIA K80 GPUs